

Error estimates for the discretization of the velocity tracking problem

Eduardo Casas^{*1}, Konstantinos Chrysafinos²

¹ Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain. e-mail: eduardo.casas@unican.es

² Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Zografou Campus, Athens 15780, Greece. e-mail: chrysafinos@math.ntua.gr

The date of receipt and acceptance will be inserted by the editor

Summary In this paper we are continuing our work [6], concerning a-priori error estimates for the velocity tracking of two-dimensional evolutionary Navier-Stokes flows. The controls are of distributed type, and subject to point-wise control constraints. The discretization scheme of the state and adjoint equations is based on a discontinuous time-stepping scheme (in time) combined with conforming finite elements (in space) for the velocity and pressure. Provided that the time and space discretization parameters, τ and h respectively, satisfy $\tau \leq Ch^2$, error estimates of order $\mathcal{O}(h^2)$ and $\mathcal{O}(h^{\frac{3}{2}-\frac{2}{p}})$ with $p > 3$ depending on the regularity of the target and the initial velocity, are proved for the difference between the locally optimal controls and their discrete approximations, when the controls are discretized by the variational discretization approach and by using piecewise-linear functions in space respectively. Both results are based on new duality arguments for the evolutionary Navier-Stokes equations.

* This author was partially supported by the Spanish Ministerio de Economía y Competitividad under project MTM2011-22711

1 Introduction

In this paper we are continuing our work of [6] regarding the approximation of the following velocity tracking problem:

$$(P) \quad \begin{cases} \min J(\mathbf{u}) \\ \mathbf{u} \in \mathcal{U}_{ad} \end{cases}$$

where

$$J(\mathbf{u}) = \frac{1}{2} \int_0^T \int_{\Omega} |\mathbf{y}_{\mathbf{u}}(t, x) - \mathbf{y}_d(t, x)|^2 dx dt + \frac{\gamma}{2} \int_{\Omega} |\mathbf{y}_{\mathbf{u}}(T, x) - \mathbf{y}_{\Omega}(x)|^2 dx \\ + \frac{\lambda}{2} \int_0^T \int_{\Omega} |\mathbf{u}(t, x)|^2 dx dt.$$

Here $\mathbf{y}_{\mathbf{u}}$ denotes the solution of the 2d evolution Navier-Stokes equations

$$\begin{cases} \mathbf{y}_t - \nu \Delta \mathbf{y} + (\mathbf{y} \cdot \nabla) \mathbf{y} + \nabla p = \mathbf{f} + \mathbf{u} & \text{in } \Omega_T = (0, T) \times \Omega, \\ \operatorname{div} \mathbf{y} = 0 & \text{in } \Omega_T, \quad \mathbf{y}(0) = \mathbf{y}_0 & \text{in } \Omega, \\ \mathbf{y} = 0 & \text{on } \Sigma_T = (0, T) \times \Gamma, \end{cases} \quad (1.1)$$

and \mathcal{U}_{ad} is the set of feasible controls, defined for $-\infty \leq \alpha_j < \beta_j \leq +\infty$, $j = 1, 2$, by

$$\mathcal{U}_{ad} = \{\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega)) : \alpha_j \leq u_j \leq \beta_j \text{ a.e. in } \Omega_T, j = 1, 2\}.$$

The scope of the velocity tracking problem is to influence and eventually drive the velocity vector field to a given target field, by using a control function of distributed type. In our setting, the control function satisfies certain constraints and $\lambda > 0$ is a penalty parameter, which is typically small compared to the actual size of the data. The inclusion of the terminal term is motivated by the need to obtain effective approximations at the end point of the time interval. We refer the reader to [21] for an related discussions, references regarding the computational significance of such optimal control problem. The analysis of the above control problem is well understood, (see e.g. [1, 5, 21, 26, 40, 44] and references within), where various aspects, including first and second order necessary conditions are developed and analyzed.

To the contrary, numerical analysis of such optimal control problems is quite limited. This is due to the fact that the restricted regularity of solutions of the evolutionary Navier-Stokes equations, as well as the divergence free condition, and the convective nature of the adjoint equation of the first order necessary condition, pose significant difficulties when analyzing numerical schemes. Standard

techniques developed for the numerical analysis of the uncontrolled Navier-Stokes equations can not be directly applied in the optimal control setting. Furthermore, the presence of control constraints, create many additional difficulties and hence require special techniques involving both first and second order necessary and sufficient conditions.

Our work further analyzes the numerical scheme proposed in [6] which is based on the discontinuous time-stepping Galerkin scheme for the piecewise constant time combined with standard conforming finite element subspaces for the discretization in space. In [6] we presented space-time error estimates of order $\mathcal{O}(h)$, under suitable regularity assumptions on the data, when the controls are discretized by piecewise constants in space and time. Two parameters τ and h are associated to the numerical scheme (here τ and h , indicating the size of the grids in time and space) and they were needed to satisfy the usual assumption $\tau \leq Ch^2$ in order to prove that the discrete equation has a unique solution, and our estimate was optimal in $L^2(0, T; \mathbf{H}^1(\Omega))$ norms for the state and adjoint. We emphasize that if we discretize the state equation only in time, not in space, then we cannot prove uniqueness of a solution for the resulting elliptic system. Indeed, this discrete elliptic system is very close to the stationary Navier-Stokes system, for which there is no a uniqueness result. Therefore, it is not surprising that the discretization parameter τ is needed to be small compared with h if we want to prove the uniqueness of a solution for the full discrete system. The key idea of [6] was to utilize ideas from [8] developed for the stationary Navier-Stokes, together with a detailed error analysis of the uncontrolled state and adjoint equations of the underlying scheme.

Here, we also make use of this condition to prove improved error estimates in two distinct cases. First, we prove optimal estimates of order $\mathcal{O}(h^2)$ in the $L^2(0, T; \mathbf{L}^2(\Omega))$ norms, when the variational discretization approach (see e.g. [25]) is being used, and estimates of order $\mathcal{O}(h^{\frac{3}{2} - \frac{2}{p}})$ for some $p > 3$, when piecewise linears (in space) are being used for the discretization of the controls. The main novel feature of our work is the development of suitable duality arguments for the state and adjoint equations, which result optimal estimate of order $\mathcal{O}(h^2)$ for the discretization of the control to state mapping, and its adjoint. Here, we emphasize that the convective nature of the adjoint equation also requires special attention. Then, we combine these estimates within the framework of [10, 8, 7] (related to nonlinear elliptic pde control constraint problems), by exploring a localization

argument and the second order condition. To our best knowledge our estimates are new.

1.1 Related results

The literature regarding numerical analysis for optimal control problems related to evolutionary Navier-Stokes equations is very limited. In [23, 24] convergence of a gradient algorithm is proven, in case of distributed controls, and of bounded distributed controls. Error estimates for the semi-discrete (in space) discretization are derived in [14] in case of distributed controls without control constraints by using a variational discretization approach. Space-time error estimates for the approximation of control problems associated to parabolic semilinear equations are presented in [35], by using both the variational discretization and the piecewise linears for the discretization of the controls. They discretize the state equation in two steps, first in time and then in space. Note that by taking advantage of the boundedness in time-space of the states they obtain error estimates for the controls without the assumption $\tau \leq Ch^2$. However, they make a strong second order condition that we do not need. Their approach is not easy to be translated to the control of Navier-Stokes systems because the non-linearity involves the gradient of the state and the boundedness of the states fails. Moreover, the discretization in time of the state equation leads to a stationary Navier-Stokes system, for which we cannot guarantee the uniqueness of a solution.

For earlier work on these schemes within the context of optimal control problems we refer the reader to [31], [32] for error estimates for an optimal control problem for the heat equation, with and without control constraints respectively, and to [12] for a convergence result for a semilinear parabolic optimal control problem. Error estimates for higher order discontinuous time-stepping schemes were presented in [11], while an analysis of second order Petrov-Galerkin Crank-Nicolson scheme and of a Crank-Nicolson scheme, for an optimal control problem for the heat equation were analyzed in [33] and [2] respectively where estimates of second-order (in time) are derived. However, the regularity assumptions on the control, state and adjoint variables are not present in the nonlinear setting of Navier-Stokes equations. For general results related to discontinuous time step schemes for linear parabolic uncontrolled pdes, we refer the reader to [15–18, 42] (see also references within). Finally in the recent work of [13], discontinuous time-stepping schemes of arbitrary order for the Navier-Stokes equations in 2d and 3d were examined.

Further results concerning the analysis and numerical analysis of the uncontrolled Navier-Stokes can be found in the classical works of [20], [27], [28], [41]. For several issues related to the analysis and numerics of optimal control problems we refer the reader to [43] (see also references within).

2 Assumptions and preliminary results

Ω is a bounded open and convex subset in \mathbb{R}^2 , Γ being its boundary. The outward unit normal vector to Γ at a point $x \in \Gamma$ is denoted by $\mathbf{n}(x)$. Given $0 < T < +\infty$, we denote $\Omega_T = (0, T) \times \Omega$ and $\Sigma_T = (0, T) \times \Gamma$. We fix the notation for Sobolev spaces: $\mathbf{H}^1(\Omega) = H^1(\Omega; \mathbb{R}^2)$, $\mathbf{H}_0^1(\Omega) = H_0^1(\Omega; \mathbb{R}^2)$, $\mathbf{H}^{-1}(\Omega) = (\mathbf{H}_0^1(\Omega))'$ and $\mathbf{W}^{s,p}(\Omega) = W^{s,p}(\Omega; \mathbb{R}^2)$ for $1 \leq p \leq \infty$ and $s > 0$. We also consider the spaces of integrable functions

$$L_0^2(\Omega) = \{w \in L^2(\Omega) : \int_{\Omega} w(x) \, dx = 0\};$$

$\mathbf{L}^p(\Omega) = L^p(\Omega; \mathbb{R}^2)$ and, for a given Banach space X , $L^p(0, T; X)$ will denote the integrable functions defined in $(0, T)$ and taking values in X endowed with the usual norm. Following Lions and Magenes [29, Vol. 1] we put

$$H^{2,1}(\Omega_T) = \left\{ y \in L^2(\Omega_T) : \frac{\partial y}{\partial x_i}, \frac{\partial^2 y}{\partial x_i x_j}, \frac{\partial y}{\partial t} \in L^2(\Omega_T), 1 \leq i, j \leq 2 \right\}$$

equipped with the standard norm. In [29, Vol. 1] it is proved that every element of $H^{2,1}(\Omega_T)$, after a modification over a zero measure set, is a continuous function from $[0, T] \rightarrow H^1(\Omega)$. We also set $\mathbf{H}^{2,1}(\Omega_T) = H^{2,1}(\Omega_T) \times H^{2,1}(\Omega_T)$.

We introduce the usual spaces of divergence-free vector fields:

$$\begin{aligned} \mathbf{Y} &= \{\mathbf{y} \in \mathbf{H}_0^1(\Omega) : \operatorname{div} \mathbf{y} = 0 \text{ in } \Omega\}, \\ \mathbf{H} &= \{\mathbf{y} \in \mathbf{L}^2(\Omega) : \operatorname{div} \mathbf{y} = 0 \text{ in } \Omega \text{ and } \mathbf{y} \cdot \mathbf{n} = 0 \text{ on } \Gamma\}. \end{aligned}$$

Along this paper, we will assume that $\mathbf{f}, \mathbf{y}_d \in L^2(0, T; \mathbf{L}^2(\Omega))$ and $\mathbf{y}_0, \mathbf{y}_\Omega \in \mathbf{Y}$. A solution of (1.1) will be sought in the space $\mathbf{W}(0, T) = \{\mathbf{y} \in L^2(0, T; \mathbf{Y}) : \mathbf{y}_t \in L^2(0, T; \mathbf{Y}^*)\}$. It is well known that $\mathbf{W}(0, T) \subset C_w([0, T], \mathbf{H})$, where $C_w([0, T], \mathbf{H})$ is the space of weakly continuous functions $\mathbf{y} : [0, T] \rightarrow \mathbf{H}$.

Let us introduce the weak formulation of (1.1). To this end we define the bilinear and trilinear forms $a : \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$ and

$c : \mathbf{L}^4(\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \longrightarrow \mathbb{R}$ by

$$\begin{aligned} a(\mathbf{y}, \mathbf{z}) &= \nu \int_{\Omega} (\nabla \mathbf{y} : \nabla \mathbf{z}) \, dx = \nu \sum_{i,j=1}^2 \int_{\Omega} \partial_{x_i} y_j \partial_{x_i} z_j \, dx \\ c(\mathbf{y}, \mathbf{z}, \mathbf{w}) &= \frac{1}{2} [\hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) - \hat{c}(\mathbf{y}, \mathbf{w}, \mathbf{z})] \\ \text{with } \hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) &= \sum_{i,j=1}^2 \int_{\Omega} \mathbf{y}_j \left(\frac{\partial \mathbf{z}_i}{\partial x_j} \right) \mathbf{w}_i \, dx. \end{aligned}$$

Now, we seek $\mathbf{y} \in \mathbf{W}(0, T)$ such that for a.e. $t \in (0, T)$,

$$\begin{cases} (\mathbf{y}_t, \mathbf{w}) + a(\mathbf{y}, \mathbf{w}) + c(\mathbf{y}, \mathbf{y}, \mathbf{w}) = (\mathbf{f} + \mathbf{u}, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y} \\ \mathbf{y}(0) = \mathbf{y}_0. \end{cases} \quad (2.1)$$

Above (\cdot, \cdot) denotes the scalar product in $\mathbf{L}^2(\Omega)$. This notation will be frequently used along the paper and $\|\cdot\|$ will denote the associated norm. Any other norm will be indicated by a subscript.

Equation (2.1) has a unique solution in $\mathbf{W}(0, T)$. Once the velocity \mathbf{y} is obtained, then the existence of a pressure $p \in \mathcal{D}(\Omega_T)$ is proved in such a way that the first equation of (1.1) holds in a distribution sense. Thanks to the regularity assumed on \mathbf{f} , \mathbf{y}_0 and Ω , then some extra regularity is proved for (\mathbf{y}, p) . Indeed, we have that $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ and $p \in L^2(0, T; H^1(\Omega))$, the pressure being unique up to an additive constant; see, for instance, Ladyzhenskaya [27], Lions [28], Temam [41].

The next properties of the trilinear form c will be used later. The proof can be found in many books; see [27], [28] or [41].

Lemma 1 *The trilinear form c satisfies*

$$\begin{aligned} c(\mathbf{y}, \mathbf{w}, \mathbf{z}) &= \hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) = -\hat{c}(\mathbf{y}, \mathbf{w}, \mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{Y} \text{ and } \forall \mathbf{z}, \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ c(\mathbf{y}, \mathbf{z}, \mathbf{w}) &= -c(\mathbf{y}, \mathbf{w}, \mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{z}, \mathbf{w} \in \mathbf{H}^1(\Omega), \\ c(\mathbf{y}, \mathbf{w}, \mathbf{w}) &= 0 \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{w} \in \mathbf{H}^1(\Omega). \end{aligned}$$

Moreover, the following inequalities hold

$$\begin{aligned} |c(\mathbf{y}, \mathbf{z}, \mathbf{w})| &\leq \|\mathbf{y}\|_{\mathbf{L}^p(\Omega)} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^q(\Omega)}, \quad (1/p) + (1/q) = (1/2), \\ |c(\mathbf{y}, \mathbf{z}, \mathbf{w})| &\leq \|\mathbf{y}\|_{\mathbf{L}^4(\Omega)} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^4(\Omega)}. \end{aligned}$$

By using the interpolation inequality

$$\|z\|_{L^4(\Omega)} \leq 2^{1/4} \|z\|_{L^2(\Omega)}^{1/2} \|\nabla z\|_{L^2(\Omega)}^{1/2} \quad \forall z \in H_0^1(\Omega), \quad (2.2)$$

(see [41, Lema 3.3, page 91]) we obtain $\forall \mathbf{y}, \mathbf{w} \in \mathbf{H}_0^1(\Omega)$ and $\forall \mathbf{z} \in \mathbf{H}^1(\Omega)$

$$|c(\mathbf{y}, \mathbf{z}, \mathbf{w})| \leq C \|\mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2}. \quad (2.3)$$

Returning back to the control problem (P), we will assume

$$\lambda > 0, \quad \gamma \geq 0, \quad \mathbf{y}_d \in L^2(0, T; \mathbf{L}^2(\Omega)) \quad \text{and} \quad \mathbf{y}_\Omega \in Y. \quad (2.4)$$

Since the mapping $G : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$, associating to each control \mathbf{u} the corresponding state $G(\mathbf{u}) = \mathbf{y}_\mathbf{u}$ solution of (2.1), is well defined and continuous, then the cost functional $J : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbb{R}$ is also well defined and continuous.

The rest of the session is devoted to the first and second order optimality conditions for local solutions. The reader is referred to [6] for a detailed analysis and proofs. By standard arguments it is easy to show that problem (P) has at least one global solution. Since (P) is not convex we can also have local solutions. The following theorems state for the differentiability of the mappings G and J , respectively.

Theorem 1 (Casas [5]) *The mapping G is of class C^∞ . Moreover, for any $\mathbf{u}, \mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$, if we denote $\mathbf{y}_\mathbf{u} = G(\mathbf{u})$, $\mathbf{z}_\mathbf{v} = G'(\mathbf{u})\mathbf{v}$ and $\mathbf{z}_{\mathbf{v}\mathbf{v}} = G''(\mathbf{u})\mathbf{v}^2$, then $\mathbf{z}_\mathbf{v}$ and $\mathbf{z}_{\mathbf{v}\mathbf{v}}$ are the unique solutions of the following equations, $\forall \mathbf{w} \in \mathbf{Y}$*

$$\begin{cases} (\mathbf{z}_{\mathbf{v},t}, \mathbf{w}) + a(\mathbf{z}_\mathbf{v}, \mathbf{w}) + c(\mathbf{z}_\mathbf{v}, \mathbf{y}_\mathbf{u}, \mathbf{w}) + c(\mathbf{y}_\mathbf{u}, \mathbf{z}_\mathbf{v}, \mathbf{w}) = (\mathbf{v}, \mathbf{w}), \\ \mathbf{z}_\mathbf{v}(0) = 0, \end{cases} \quad (2.5)$$

$$\begin{cases} (\mathbf{z}_{\mathbf{v}\mathbf{v},t}, \mathbf{w}) + a(\mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{w}) + c(\mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{y}_\mathbf{u}, \mathbf{w}) + c(\mathbf{y}_\mathbf{u}, \mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{w}) \\ \quad + 2c(\mathbf{z}_\mathbf{v}, \mathbf{z}_\mathbf{v}, \mathbf{w}) = 0, \\ \mathbf{z}_{\mathbf{v}\mathbf{v}}(0) = 0. \end{cases} \quad (2.6)$$

Theorem 2 (Casas and Chrysaftinos [6]) *The objective functional $J : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbb{R}$ is of class C^∞ and for every $\mathbf{u}, \mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$ we have*

$$J'(\mathbf{u})\mathbf{v} = \int_0^T \int_\Omega (\varphi_\mathbf{u} + \lambda \mathbf{u}) \mathbf{v} \, dx dt, \quad (2.7)$$

$$\begin{aligned} J''(\mathbf{u})\mathbf{v}^2 &= \int_0^T \int_\Omega (|\mathbf{z}_\mathbf{v}|^2 - 2(\mathbf{z}_\mathbf{v} \cdot \nabla) \mathbf{z}_\mathbf{v} \varphi_\mathbf{u}) \, dx dt \\ &\quad + \gamma \int_\Omega |\mathbf{z}_\mathbf{v}(T)|^2 \, dx + \lambda \int_0^T \int_\Omega |\mathbf{v}|^2 \, dx dt, \end{aligned} \quad (2.8)$$

where $\mathbf{z}_v = G'(\mathbf{u})\mathbf{v}$ is the solution of (2.5) and $\varphi_{\mathbf{u}} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ is the unique element satisfying

$$\begin{cases} -(\varphi_{\mathbf{u},t}, \mathbf{w}) + a(\varphi_{\mathbf{u}}, \mathbf{w}) + c(\mathbf{w}, \mathbf{y}_{\mathbf{u}}, \varphi_{\mathbf{u}}) + c(\mathbf{y}_{\mathbf{u}}, \mathbf{w}, \varphi_{\mathbf{u}}) \\ = (\mathbf{y}_{\mathbf{u}} - \mathbf{y}_d, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \varphi_{\mathbf{u}}(T) = \gamma(\mathbf{y}(T) - \mathbf{y}_{\Omega}). \end{cases} \quad (2.9)$$

From this theorem we can deduce the first order optimality conditions (see e.g. [6, Theorem 3.3]).

Theorem 3 *Let us assume that $\bar{\mathbf{u}}$ is a local solution of problem (P), then there exist $\bar{\mathbf{y}}$ and $\bar{\varphi}$ belonging to $\mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ such that*

$$\begin{cases} (\bar{\mathbf{y}}_t, \mathbf{w}) + a(\bar{\mathbf{y}}, \mathbf{w}) + c(\bar{\mathbf{y}}, \bar{\mathbf{y}}, \mathbf{w}) = (\mathbf{f} + \bar{\mathbf{u}}, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \bar{\mathbf{y}}(0) = \mathbf{y}_0, \end{cases} \quad (2.10)$$

$$\begin{cases} -(\bar{\varphi}_t, \mathbf{w}) + a(\bar{\varphi}, \mathbf{w}) + c(\mathbf{w}, \bar{\mathbf{y}}, \bar{\varphi}) + c(\bar{\mathbf{y}}, \mathbf{w}, \bar{\varphi}) \\ = (\bar{\mathbf{y}}_{\mathbf{u}} - \mathbf{y}_d, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \bar{\varphi}(T) = \gamma(\bar{\mathbf{y}}(T) - \mathbf{y}_{\Omega}), \end{cases} \quad (2.11)$$

$$\int_0^T \int_{\Omega} (\bar{\varphi} + \lambda \bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \, dx dt \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad}. \quad (2.12)$$

Moreover, the regularity property $\bar{\mathbf{u}} \in \mathbf{H}^1(\Omega_T) \cap C([0, T], \mathbf{H}^1(\Omega)) \cap L^2(0, T; \mathbf{W}^{1,p}(\Omega))$ holds for all $1 \leq p < +\infty$.

From relation (2.12), we get the usual projection formula for $j = 1, 2$

$$\bar{u}_j(t, x) = \text{Proj}_{[\alpha_j, \beta_j]} \left(-\frac{1}{\lambda} \bar{\varphi}_j(t, x) \right) \quad \text{for a.a. } (t, x) \in \Omega_T. \quad (2.13)$$

To write the second order conditions we need to define the cone of critical directions. To this end, let us introduce the function

$$\bar{\mathbf{d}} = \bar{\varphi} + \lambda \bar{\mathbf{u}}. \quad (2.14)$$

Now we set

$$\mathcal{C}_{\bar{\mathbf{u}}} = \{ \mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega)) : \mathbf{v} \text{ satisfies (2.16) - (2.18)} \}, \quad (2.15)$$

$$v_j(t, x) \geq 0 \text{ if } -\infty < \alpha_j = \bar{u}_j(t, x), \quad (2.16)$$

$$v_j(t, x) \leq 0 \text{ if } \bar{u}_j(t, x) = \beta_j < +\infty, \quad j = 1, 2, \quad (2.17)$$

$$v_j(t, x) = 0 \text{ if } \bar{d}_j(t, x) \neq 0. \quad (2.18)$$

Let us notice that

$$J'(\bar{\mathbf{u}})\mathbf{v} = \int_0^T \int_{\Omega} \bar{\mathbf{d}}(t, x) \cdot \mathbf{v}(t, x) \, dx dt, \quad (2.19)$$

$$\bar{\mathbf{d}}(t, x) \cdot \mathbf{v}(t, x) = 0 \text{ for a.a. } (t, x) \in \Omega_T \text{ and } \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}.$$

We also deduce as usual from (2.12), for almost all $(t, x) \in \Omega_T$ and $j = 1, 2$,

$$\begin{cases} \bar{u}_j(t, x) = \alpha_j & \Rightarrow \bar{d}_j(t, x) \geq 0, \\ \bar{u}_j(t, x) = \beta_j & \Rightarrow \bar{d}_j(t, x) \leq 0, \\ \alpha_j < \bar{u}_j(t, x) < \beta_j & \Rightarrow \bar{d}_j(t, x) = 0, \\ \bar{d}_j(t, x) > 0 & \Rightarrow \bar{u}_j(t, x) = \alpha_j, \\ \bar{d}_j(t, x) < 0 & \Rightarrow \bar{u}_j(t, x) = \beta_j. \end{cases} \quad (2.20)$$

Theorem 4 *Let $\bar{\mathbf{u}}$ be a local solution of problem (P), then $J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq 0 \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. Conversely, let us assume that $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ satisfies*

$$J'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad}, \quad (2.21)$$

$$J''(\bar{\mathbf{u}})\mathbf{v}^2 > 0 \quad \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}} \setminus \{0\}, \quad (2.22)$$

then there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$J(\bar{\mathbf{u}}) + \frac{\delta}{2} \|\mathbf{u} - \bar{\mathbf{u}}\|_{L^2(0, T; \mathbf{L}^2(\Omega))}^2 \leq J(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{U}_{ad} \cap B_{\varepsilon}(\bar{\mathbf{u}}), \quad (2.23)$$

where $B_{\varepsilon}(\bar{\mathbf{u}})$ is the $L^2(0, T; \mathbf{L}^2(\Omega))$ -ball of center $\bar{\mathbf{u}}$ and radius ε .

Finally, we point out that the sufficient second order condition (2.22) is equivalent to the following:

$$\exists \delta > 0 \quad \text{such that} \quad J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq \delta \|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}. \quad (2.24)$$

3 Approximation of the state and adjoint-state equations

We consider a family of triangulations $\{\mathcal{K}_h\}_{h>0}$ of $\bar{\Omega}$, defined in the standard way. To each element $K \in \mathcal{K}_h$, we associate two parameters h_K and ϱ_K , where h_K denotes the diameter of the set K and ϱ_K is the diameter of the largest ball contained in K . Define the size of the mesh by $h = \max_{K \in \mathcal{K}_h} h_K$. We also assume that the standard regularity assumptions on the triangulation:

(i) – There exist two positive constants $\varrho_{\mathcal{K}}$ and $\delta_{\mathcal{K}}$ such that $\frac{h_K}{\varrho_K} \leq \varrho_{\mathcal{K}}$ and $\frac{h}{h_K} \leq \delta_{\mathcal{K}} \forall K \in \mathcal{K}_h$ and $\forall h > 0$.

(ii) – Define $\overline{\Omega}_h = \cup_{K \in \mathcal{K}_h} K$, and let Ω_h and Γ_h denote its interior and its boundary, respectively. We assume that the vertices of \mathcal{K}_h placed on the boundary Γ_h are points of Γ .

Since Ω is convex, from the last assumption we have that Ω_h is also convex. Moreover, we know that

$$|\Omega \setminus \Omega_h| \leq Ch^2; \quad (3.1)$$

see, for instance, [36, estimate (5.2.19)].

On the mesh \mathcal{K}_h we consider two finite dimensional spaces $\mathbf{Z}_h \subset \mathbf{H}_0^1(\Omega)$ and $Q_h \subset L_0^2(\Omega)$ formed by piecewise polynomials in Ω_h and vanishing in $\Omega \setminus \Omega_h$. We make the following assumptions on these spaces.

(A1) If $\mathbf{z} \in \mathbf{H}^{1+l}(\Omega) \cap \mathbf{H}_0^1(\Omega)$, then

$$\inf_{\mathbf{z}_h \in \mathbf{Z}_h} \|\mathbf{z} - \mathbf{z}_h\|_{\mathbf{H}^s(\Omega_h)} \leq Ch^{l+1-s} \|\mathbf{z}\|_{\mathbf{H}^{1+l}(\Omega)}, \quad \text{for } 0 \leq l \leq 1 \text{ and } s = 0, 1. \quad (3.2)$$

(A2) If $q \in H^l(\Omega) \cap L_0^2(\Omega)$, then

$$\inf_{q_h \in Q_h} \|q - q_h\|_{L^2(\Omega_h)} \leq Ch \|q\|_{H^1(\Omega)}. \quad (3.3)$$

(A3) The subspaces \mathbf{Z}_h and Q_h satisfy the inf-sup condition: $\exists c > 0$ such that

$$\inf_{q_h \in Q_h} \sup_{\mathbf{z}_h \in \mathbf{Z}_h} \frac{b(\mathbf{z}_h, q_h)}{\|\mathbf{z}_h\|_{\mathbf{H}^1(\Omega_h)} \|q_h\|_{L^2(\Omega_h)}} \geq c, \quad (3.4)$$

where $b : \mathbf{H}^1(\Omega) \times L^2(\Omega) \rightarrow \mathbb{R}$ is defined by

$$b(\mathbf{z}, q) = \int_{\Omega} q(x) \operatorname{div} \mathbf{z}(x) \, dx.$$

These assumptions are satisfied by the usual finite elements considered in the discretization of Navier-Stokes equations: "Taylor-Hood", P1-Bubble finite element, and some others; see [20, Chapter 2].

We also consider a subspace \mathbf{Y}_h of \mathbf{Z}_h defined by

$$\mathbf{Y}_h = \{\mathbf{y}_h \in \mathbf{Z}_h : b(\mathbf{y}_h, q_h) = 0 \, \forall q_h \in Q_h\}.$$

We proceed now with the discretization in time. Let us consider a grid of points $0 = t_0 < t_1 < \dots < t_{N_\tau} = T$. We denote $\tau_n = t_n - t_{n-1}$. We make the following assumption

$$\exists \varrho_0 > 0 \text{ s.t. } \tau = \max_{1 \leq n \leq N_\tau} \tau_n < \varrho_0 \tau_n \quad \forall 1 \leq n \leq N_\tau \text{ and } \forall \tau > 0. \quad (3.5)$$

Given a triangulation \mathcal{K}_h of Ω and a grid of points $\{t_n\}_{n=0}^{N_\tau}$ of $[0, T]$, we set $\sigma = (\tau, h)$. Finally, we consider the following spaces

$$\begin{aligned} \mathcal{Y}_\sigma &= \{\mathbf{y}_\sigma \in L^2(0, T; \mathbf{Y}_h) : \mathbf{y}_\sigma|_{(t_{n-1}, t_n)} \in \mathbf{Y}_h \text{ for } 1 \leq n \leq N_\tau\}, \\ \mathcal{Q}_\sigma &= \{q_\sigma \in L^2(0, T; Q_h) : q_\sigma|_{(t_{n-1}, t_n)} \in Q_h \text{ for } 1 \leq n \leq N_\tau\}. \end{aligned}$$

We have that the functions of \mathcal{Y}_σ and \mathcal{Q}_σ are piecewise constant in time. The elements of \mathcal{Y}_σ can be written in the form

$$\mathbf{y}_\sigma = \sum_{n=1}^{N_\tau} \mathbf{y}_{n,h} \chi_n, \quad \text{with } \mathbf{y}_{n,h} \in \mathbf{Y}_h, \quad (3.6)$$

where χ_n is the characteristic function of (t_{n-1}, t_n) . For every discrete state \mathbf{y}_σ we will fix $\mathbf{y}_\sigma(t_n) = \mathbf{y}_{n,h}$, so that \mathbf{y}_σ is continuous on the left. In particular, we have $\mathbf{y}_\sigma(T) = \mathbf{y}_\sigma(t_{N_\tau}) = \mathbf{y}_{N_\tau,h}$.

3.1 The discrete state equation

To define the discrete control problem we have to consider the numerical discretization of the state equation (1.1) or equivalently (2.1). We achieve this goal by using a discontinuous time-stepping Galerkin method, with piecewise constants in time and conforming finite element spaces in space. For any $\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$ the discrete state equation is given by

$$\begin{cases} \text{For } n = 1, \dots, N_\tau, \\ \left(\frac{\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\mathbf{y}_{n,h}, \mathbf{w}_h) + c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) \\ = (\mathbf{f}_n + \mathbf{u}_n, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \mathbf{y}_{0,h} = \mathbf{y}_{0h}, \end{cases} \quad (3.7)$$

where

$$(\mathbf{f}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{f}(t), \mathbf{w}_h) dt, \quad (\mathbf{u}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{u}(t), \mathbf{w}_h) dt, \quad (3.8)$$

$$\mathbf{y}_{0h} \in \mathbf{Y}_h \text{ with } \|\mathbf{y}_0 - \mathbf{y}_{0h}\|_{\mathbf{L}^2(\Omega_h)} \leq Ch, \text{ and } \|\mathbf{y}_{0h}\|_{\mathbf{H}^1(\Omega_h)} \leq C. \quad (3.9)$$

Let us consider the projection operator $P_h : \mathbf{L}^2(\Omega) \rightarrow \mathbf{Y}_h$ given by

$$(P_h \mathbf{y}, \mathbf{w}_h) = (\mathbf{y}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h.$$

We also define $P_\sigma : C([0, T], \mathbf{L}^2(\Omega)) \rightarrow \mathcal{Y}_\sigma$ by $(P_\sigma \mathbf{y})_{n,h} = P_h \mathbf{y}(t_n)$ for $1 \leq n \leq N_\tau$. This is the standard projection associated to the discontinuous time-stepping scheme of the lowest order in time. The next estimate was proven in [6, Lemma 4.2].

Lemma 2 *There exists a constant $C > 0$ independent of σ such that for every $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$ the following estimates hold*

$$\begin{aligned} & \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \\ & \leq C \left\{ \tau \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h^2 \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\}, \end{aligned} \quad (3.10)$$

$$\begin{aligned} & \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \\ & \leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\}. \end{aligned} \quad (3.11)$$

It well known that the discrete equation (3.7) has at least one solution. Concerning uniqueness and error estimates under the prescribed regularity assumptions, the following results was proven in [6, Theorem 4.7].

Theorem 5 *Given $\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$, let us denote the solution of (2.1) by $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$, and let $\mathbf{y}_\sigma \in \mathcal{Y}_\sigma$ be any solution of (3.7). Then, there exists a constant $C > 0$ independent of \mathbf{u} , \mathbf{y} and σ such that*

$$\begin{aligned} & \max_{1 \leq n \leq N_\tau} \|\mathbf{y}(t_n) - \mathbf{y}_\sigma(t_n)\| + \|\mathbf{y} - \mathbf{y}_\sigma\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \\ & \leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}. \end{aligned} \quad (3.12)$$

$$\begin{aligned} & \|\mathbf{y} - \mathbf{y}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} \leq C \left\{ \left(\frac{\tau}{h} + \sqrt{\tau} \right) \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} \right. \\ & \left. + h \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}. \end{aligned} \quad (3.13)$$

Moreover, if there exists a constant $C_0 > 0$ such that $\tau \leq C_0 h^2$ for every $\sigma = (\tau, h)$, then $\{\mathbf{y}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ and (3.7) has a unique solution.

By using the above results, and a suitable duality argument we deduce the improved estimate on $L^2[0, T; \mathbf{L}^2(\Omega)]$ for state variable.

Theorem 6 *Under the notation and assumptions of theorem 5 and including the hypothesis $\tau \leq C_0 h^2$ for some constant C_0 , the following estimate holds:*

$$\|\mathbf{y} - \mathbf{y}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \leq C h^2, \quad (3.14)$$

where C is independent of σ , \mathbf{f} and \mathbf{u} .

Proof Let us define $\mathbf{e} = \mathbf{y} - \mathbf{y}_\sigma = (\mathbf{y} - \hat{\mathbf{y}}_\sigma) + (\hat{\mathbf{y}}_\sigma - \mathbf{y}_\sigma) = \hat{\mathbf{e}} + \mathbf{e}_\sigma$, where $\hat{\mathbf{y}}_\sigma$ is the solution of the following discrete equation:

$$\begin{cases} \text{For } n = 1, \dots, N_\tau, \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\hat{\mathbf{y}}_{n,h}, \mathbf{w}_h) = (\hat{\mathbf{f}}_n, \mathbf{w}_h), \\ \hat{\mathbf{y}}_{0,h} = \mathbf{y}_{0h}, \end{cases} \quad (3.15)$$

where $(\hat{\mathbf{f}}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \{a(\mathbf{y}(t), \mathbf{w}_h) + (\mathbf{y}'(t), \mathbf{w}_h)\} dt$. Then, (3.15) has a unique solution $\hat{\mathbf{y}}_\sigma \in \mathcal{Y}_\sigma$. Moreover, [6, Lemma 4.6] states that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$, and in addition there exists a constant $C > 0$ independent of σ such that the following estimates hold:

$$\max_{1 \leq n \leq N_\tau} \|\mathbf{y}(t_n) - \hat{\mathbf{y}}_\sigma(t_n)\| + \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \quad (3.16)$$

$$\leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}, \quad (3.17)$$

$$\|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} \leq C \left\{ \left(\frac{\tau}{h} + \sqrt{\tau} \right) \|\mathbf{y}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}.$$

The proof consists of two parts. First, we employ a duality argument to deduce the inequality,

$$\|\mathbf{e}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \leq C(\|\hat{\mathbf{e}}\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h^2), \quad (3.18)$$

with C independent of σ . Second, we estimate the first term on the right hand side, and then inserting (3.16) into (3.18) we obtain the increased rate (3.14).

Step 1-Proof of (3.18). We begin by defining a dual problem

$$\begin{cases} -(\phi_t, \mathbf{w}) + a(\phi, \mathbf{w}) + c(\mathbf{w}, \mathbf{y}_\sigma, \phi) + c(\mathbf{y}, \mathbf{w}, \phi) \\ = (\mathbf{e}_\sigma, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \phi(T) = 0. \end{cases} \quad (3.19)$$

Note that ϕ has similar convective structure to the adjoint equation, but it is different from the solution $\varphi_{\mathbf{u}}$ of (2.9). Recall that according to Theorem 5, the sequence $\{\mathbf{y}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega))$ and $\{\mathbf{e}_\sigma\}_\sigma$ is also bounded in $L^2(0, T; \mathbf{L}^2(\Omega))$. Well known regularity theorems state that $\phi \in \mathbf{H}^{2,1}(\Omega_T) \cap L^\infty(0, T; \mathbf{Y})$ and its norm is estimated by $\|\mathbf{e}_\sigma\|_{L^2(\Omega_T)}$; see for instance [5]. The discrete version

(within the context of discontinuous time-stepping schemes) takes the form:

$$\begin{cases} \text{for } n = N_\tau, \dots, 1, \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ (\phi_{n,h} - \phi_{n+1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} \{a(\phi_{n,h}, \mathbf{w}_h) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \phi_{n,h}) \\ + c(\mathbf{y}(t), \mathbf{w}_h, \phi_{n,h})\} dt = \int_{t_{n-1}}^{t_n} (\mathbf{e}_{n,h}, \mathbf{w}_h) dt, \\ \phi_{N_\tau+1,h} = 0. \end{cases} \quad (3.20)$$

Note that unlike the discrete states \mathbf{y}_σ , we will set for the discrete adjoint states $\phi_\sigma(t_{n-1}) = \phi_{n,h}$ for every $1 \leq n \leq N_\tau$. Now, [6, Theorem 4.11] implies that

$$\|\phi - \phi_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq Ch \|\mathbf{e}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}. \quad (3.21)$$

Setting $\mathbf{w}_h = \mathbf{e}_{n,h} = \mathbf{e}_\sigma(t_n)$ in (3.20), we obtain,

$$\begin{aligned} \int_{t_{n-1}}^{t_n} \|\mathbf{e}_{n,h}\|_{\mathbf{L}^2(\Omega_h)}^2 dt &= (\phi_{n,h} - \phi_{n+1,h}, \mathbf{e}_{n,h}) \\ &+ \int_{t_{n-1}}^{t_n} \{a(\phi_{n,h}, \mathbf{e}_{n,h}) + c(\mathbf{e}_{n,h}, \mathbf{y}_{n,h}, \phi_{n,h}) + c(\mathbf{y}(t), \mathbf{e}_{n,h}, \phi_{n,h})\} dt. \end{aligned} \quad (3.22)$$

From the definition (3.15) of $\hat{\mathbf{y}}_\sigma$, and taking $\mathbf{w} = \mathbf{w}_h$ in (2.1), we obtain for every $1 \leq n \leq N_\tau$

$$(\hat{\mathbf{e}}(t_n) - \hat{\mathbf{e}}(t_{n-1}), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}(t), \mathbf{w}_h) dt = 0 \quad \forall \mathbf{w}_h \in \mathbf{Y}_h. \quad (3.23)$$

Subtracting (3.7) from (2.1) with $\mathbf{w} \equiv \mathbf{w}_h$, the orthogonality condition for the discrete state equation takes the form

$$\begin{aligned} (\mathbf{e}(t_n) - \mathbf{e}(t_{n-1}), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\mathbf{e}(t), \mathbf{w}_h) dt \\ = \int_{t_{n-1}}^{t_n} \{c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) - c(\mathbf{y}(t), \mathbf{y}(t), \mathbf{w}_h)\} dt. \end{aligned} \quad (3.24)$$

Using the decomposition $\mathbf{e} = \hat{\mathbf{e}} + \mathbf{e}_\sigma$ into the previous relation and invoking (3.23) with $\mathbf{w}_h \equiv \phi_{n,h}$ it follows,

$$\begin{aligned} (\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}, \phi_{n,h}) + \int_{t_{n-1}}^{t_n} a(\mathbf{e}_{n,h}, \phi_{n,h}) dt \\ = \int_{t_{n-1}}^{t_n} \{c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \phi_{n,h}) - c(\mathbf{y}(t), \mathbf{y}(t), \phi_{n,h})\} dt. \end{aligned} \quad (3.25)$$

For the last two nonlinear terms, using the definition of \mathbf{e} and the decomposition $\mathbf{e} = \hat{\mathbf{e}} + \mathbf{e}_\sigma$, we obtain:

$$\begin{aligned} & c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \phi_{n,h}) - c(\mathbf{y}(t), \mathbf{y}(t), \phi_{n,h}) \\ &= -c(\mathbf{e}(t), \mathbf{y}_{n,h}, \phi_{n,h}) - c(\mathbf{y}(t), \mathbf{e}(t), \phi_{n,h}) \\ &= -c(\hat{\mathbf{e}}(t), \mathbf{y}_{n,h}, \phi_{n,h}) - c(\mathbf{e}_{n,h}, \mathbf{y}_{n,h}, \phi_{n,h}) \\ &\quad -c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi_{n,h}) - c(\mathbf{y}(t), \mathbf{e}_{n,h}, \phi_{n,h}) \end{aligned}$$

Combining this identity with (3.25) we arrive to:

$$\begin{aligned} & (\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}, \phi_{n,h}) + \int_{t_{n-1}}^{t_n} \{a(\mathbf{e}_{n,h}, \phi_{n,h}) \\ &+ c(\mathbf{e}_{n,h}, \mathbf{y}_{n,h}, \phi_{n,h}) + c(\mathbf{y}(t), \mathbf{e}_{n,h}, \phi_{n,h})\} dt \\ &= - \int_{t_{n-1}}^{t_n} \{c(\hat{\mathbf{e}}(t), \mathbf{y}_{n,h}, \phi_{n,h}) + c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi_{n,h})\} dt. \end{aligned} \quad (3.26)$$

Combining now (3.22) and (3.26), we get:

$$\begin{aligned} & \int_{t_{n-1}}^{t_n} \|\mathbf{e}_{n,h}\|_{\mathbf{L}^2(\Omega_h)}^2 dt = (\phi_{n,h} - \phi_{n+1,h}, \mathbf{e}_{n,h}) - (\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}, \phi_{n,h}) \\ & \quad - \int_{t_{n-1}}^{t_n} \{c(\hat{\mathbf{e}}(t), \mathbf{y}_{n,h}, \phi_{n,h}) + c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi_{n,h})\} dt. \end{aligned}$$

Summing the above equalities, from $n = 1$ to $n = N_\tau$, and noting that $\phi_{N_\tau,h} = 0$ and $\mathbf{e}_{0,h} = 0$ we deduce that

$$\begin{aligned} & \|\mathbf{e}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 \\ &= - \int_0^T \{c(\hat{\mathbf{e}}(t), \mathbf{y}_\sigma(t), \phi_\sigma(t)) + c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi_\sigma(t))\} dt. \end{aligned} \quad (3.27)$$

Using (3.16), (3.21), along with Lemma 1 and the boundedness of $\{\mathbf{y}_\sigma\}_\sigma$ in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ stated in Theorem 5, we deduce from (3.27)

$$\begin{aligned} & \|\mathbf{e}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 = - \int_0^T \{c(\hat{\mathbf{e}}(t), \mathbf{y}_\sigma(t), \phi(t)) + c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi(t))\} dt \\ & \quad + \int_0^T \{c(\hat{\mathbf{e}}(t), \mathbf{y}_\sigma(t), \phi(t) - \phi_\sigma(t)) + c(\mathbf{y}(t), \hat{\mathbf{e}}(t), \phi(t) - \phi_\sigma(t))\} dt \\ & \leq C \|\hat{\mathbf{e}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} (\|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}\|_{L^\infty(0,T;\mathbf{H}^1(\Omega))}) \|\phi\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \\ & \quad + C (\|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}\|_{L^\infty(0,T;\mathbf{H}^1(\Omega))}) \|\hat{\mathbf{e}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \|\phi - \phi_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \end{aligned}$$

$$\leq C(\|\hat{\mathbf{e}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + h^2)\|\mathbf{e}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.$$

This completes the proof of the first step.

Step 2-Proof of the estimate for $\|\hat{\mathbf{e}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}$. We begin by considering the decomposition $\hat{\mathbf{e}} = \hat{\mathbf{e}}_p + \hat{\mathbf{e}}_\sigma$ with $\hat{\mathbf{e}}_p = \mathbf{y} - P_\sigma \mathbf{y}$ and $\hat{\mathbf{e}}_\sigma = P_\sigma \mathbf{y} - \hat{\mathbf{y}}_\sigma$. From Lemma 1 for $\tau \leq C_0 h^2$, we know that

$$\|\hat{\mathbf{e}}_p\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + h\|\hat{\mathbf{e}}_p\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq Ch^2\|\mathbf{y}\|_{\mathbf{H}^{2,1}(\Omega_T)}. \quad (3.28)$$

It remains to estimate $\hat{\mathbf{e}}_\sigma$. For the later, we proceed by defining a suitable duality argument. To this end we now introduce $\phi \in \mathbf{H}^{2,1}(\Omega_T)$, solution of,

$$\begin{cases} -(\phi_t, \mathbf{w}) + a(\phi, \mathbf{w}) = (\hat{\mathbf{e}}_\sigma, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \phi(T) = 0, \end{cases} \quad (3.29)$$

The discrete version takes the form:

$$\begin{cases} \text{for } n = N_\tau, \dots, 1, \\ (\phi_{n,h} - \phi_{n+1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\phi_{n,h}, \mathbf{w}_h) dt \\ = \int_{t_{n-1}}^{t_n} (\hat{\mathbf{e}}_{n,h}, \mathbf{w}_h) dt, \quad \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \phi_{N_\tau+1,h} = 0. \end{cases} \quad (3.30)$$

Inserting the decomposition of $\hat{\mathbf{e}} = \hat{\mathbf{e}}_p + \hat{\mathbf{e}}_\sigma$ into (3.23), and taking into account that $(\hat{\mathbf{e}}_p(t_n), \mathbf{w}_h) = (\mathbf{y}(t_n) - P_h \mathbf{y}(t_n), \mathbf{w}_h) = 0$ for every $n = 1, \dots, N_\tau$ and all $\mathbf{w}_h \in \mathbf{Y}_h$, we obtain $\forall \mathbf{w}_h \in \mathbf{Y}_h, 1 \leq n \leq N_\tau$.

$$(\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}_{n,h}, \mathbf{w}_h) dt = - \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}_p(t), \mathbf{w}_h) dt. \quad (3.31)$$

Taking $\mathbf{w}_h = \hat{\mathbf{e}}_{n,h}$ into (3.30), and $\mathbf{w}_h = \phi_{n,h}$ into (3.31), and subtracting the resulting equations, we arrive to

$$\begin{aligned} \|\hat{\mathbf{e}}_{n,h}\|_{L^2(t_{n-1}, t_n; \mathbf{L}^2(\Omega_h))}^2 &= (\phi_{n,h} - \phi_{n+1,h}, \hat{\mathbf{e}}_{n,h}) - (\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}, \phi_{n,h}) \\ &\quad - \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}_p(t), \phi_{n,h}) dt. \end{aligned}$$

Summing the above equalities from $n = 1$ to $n = N_\tau$, inserting ϕ on the right hand side, we obtain

$$\begin{aligned} &\|\hat{\mathbf{e}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 \\ &\leq - \int_0^T a(\hat{\mathbf{e}}_p(t), \phi_\sigma(t) - \phi(t)) dt + \int_0^T a(\hat{\mathbf{e}}_p(t), \phi(t)) dt. \end{aligned}$$

From the linear equations (3.29) and (3.30), working similarly to [13, Theorem 4.6], we obtain $\|\phi_\sigma - \phi\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq Ch\|\hat{\mathbf{e}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}$. Using this fact in the above inequality and (3.28), and integrating by parts in the last integral, we finally conclude

$$\begin{aligned} \|\hat{\mathbf{e}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 &\leq C(\|\hat{\mathbf{e}}_p\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\|\phi_\sigma - \phi\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ &+ \|\hat{\mathbf{e}}_p\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\|\phi\|_{L^2(0,T;\mathbf{H}^2(\Omega))}) \leq Ch^2\|\hat{\mathbf{e}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}. \end{aligned}$$

Combining Theorem 6 with [6, Corollary 4.11], we get

Corollary 1 *Assume that $\max\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}, \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\} \leq M$. Let $\mathbf{y}_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ be the solution of (2.1) and $\mathbf{y}_\sigma(\mathbf{v}) \in \mathcal{Y}_\sigma$ the solution of the discrete equation (3.7) corresponding to the control \mathbf{v} . Then, there exists a constant $C_M > 0$ such that*

$$\begin{aligned} &\|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\ &\quad + h(\|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}) \\ &\leq C_M \{h^2 + \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\} \end{aligned} \quad (3.32)$$

Moreover, if $\mathbf{u}_\sigma \in \mathcal{U}_\sigma \forall \sigma$ and $\mathbf{u}_\sigma \rightharpoonup \mathbf{u}$ weakly in $L^2(0,T;\mathbf{L}^2(\Omega))$, then

$$\begin{cases} \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \rightarrow 0, \\ \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^p(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0 \quad \forall 1 \leq p < +\infty, \\ \|\mathbf{y}_u(T) - \mathbf{y}_\sigma(\mathbf{u}_\sigma)(T)\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0. \end{cases} \quad (3.33)$$

3.2 The discrete adjoint-state equation

In this section, the assumption $\tau \leq C_0 h^2$ will be assumed. Then, associated to the discrete state equation (3.7), the cost functional J is approximated by $J_\sigma : L^2(0,T;\mathbf{L}^2(\Omega)) \rightarrow \mathbb{R}$, given by

$$\begin{aligned} J_\sigma(\mathbf{u}) &= \frac{1}{2} \int_0^T \int_{\Omega_h} |\mathbf{y}_\sigma(\mathbf{u}) - \mathbf{y}_d|^2 dx dt \\ &\quad + \frac{\gamma}{2} \int_{\Omega_h} |\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}|^2 dx + \frac{\lambda}{2} \int_0^T \int_{\Omega_h} |\mathbf{u}|^2 dx dt, \end{aligned}$$

where $\mathbf{y}_{\Omega_h} \in \mathbf{Y}_h$ is a convenient approximation of \mathbf{y}_Ω to be fixed later and $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u})$ is the discrete state corresponding to the control \mathbf{u} .

The functional J_σ is of class C^∞ , and we have a first expression of its derivative as follows

$$\begin{aligned} J'_\sigma(\mathbf{u})\mathbf{v} &= \int_0^T \int_{\Omega_h} (\mathbf{y}_\sigma - \mathbf{y}_d) \mathbf{z}_\sigma \, dx dt \\ &+ \gamma \int_{\Omega_h} (\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}) \mathbf{z}_\sigma(T) \, dx + \lambda \int_0^T \int_{\Omega_h} \mathbf{u} \mathbf{v} \, dx dt, \end{aligned}$$

where \mathbf{z}_σ is the solution of the linearized equation

$$\left\{ \begin{array}{l} \text{For } n = 1, \dots, N_\tau, \\ \left(\frac{\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\mathbf{z}_{n,h}, \mathbf{w}_h) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) \\ + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \mathbf{w}_h) \, dt \quad \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \mathbf{z}_{0,h} = 0; \end{array} \right. \quad (3.34)$$

see [6, Theorem 4.12]. By using the adjoint state equation

$$\left\{ \begin{array}{l} \text{for } n = N_\tau, \dots, 1, \quad \text{and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\varphi_{n,h} - \varphi_{n+1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\varphi_{n,h}, \mathbf{w}_h) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \varphi_{n,h}) \\ + c(\mathbf{y}_{n,h}, \mathbf{w}_h, \varphi_{n,h}) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{y}_{n,h} - \mathbf{y}_d(t), \mathbf{w}_h) \, dt, \\ \varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h}), \end{array} \right. \quad (3.35)$$

the derivative of J_σ can be expressed as

$$J'_\sigma(\mathbf{u})\mathbf{v} = \int_0^T \int_{\Omega_h} (\varphi_\sigma + \lambda \mathbf{u}) \mathbf{v} \, dx dt. \quad (3.36)$$

Observe that in the above system (3.35), first we compute $\varphi_{N_\tau,h}$ from $\varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h})$ and then we descend in n until $n = 1$. As mentioned before, unlike the discrete states \mathbf{y}_σ , we will set for the discrete adjoint states $\varphi_\sigma(t_{n-1}) = \varphi_{n,h}$ for every $1 \leq n \leq N_\tau$. Analogously to the Corollary 1, we have the following result.

Theorem 7 *Assume that $\max\{\|\mathbf{u}\|_{L^2(0,T;L^2(\Omega))}, \|\mathbf{v}\|_{L^2(0,T;L^2(\Omega))}\} \leq M$. Let $\varphi_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T]; \mathbf{Y})$ be the solution of (2.9) and $\varphi_\sigma(\mathbf{v}) \in \mathcal{Y}_\sigma$ the solution of the discrete equation (3.35) corresponding to the control \mathbf{v} . Then, there exists a constant $C_M > 0$ such that*

$$\begin{aligned} &\|\varphi_u - \varphi_\sigma(\mathbf{v})\|_{L^2(0,T;L^2(\Omega_h))} \\ &+ h(\|\varphi_u - \varphi_\sigma(\mathbf{v})\|_{L^\infty(0,T;L^2(\Omega_h))} + \|\varphi_u - \varphi_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}) \\ &\leq C_M \{h^2 + \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;L^2(\Omega))}\} \end{aligned} \quad (3.37)$$

Proof The only difference of this Theorem with respect to [6, Corollary 4.14] (see also [6, Theorem 4.13]) is the $L^2(0, T; \mathbf{L}^2(\Omega_h))$ estimate for the discretization of the adjoint state equation. Let us prove these estimates. For simplicity we denote $\boldsymbol{\varphi}$ instead of $\boldsymbol{\varphi}_{\mathbf{u}}$ and similarly for its discrete approximation. First, we define $R_\sigma : C([0, T]; \mathbf{L}^2(\Omega)) \rightarrow \mathcal{Y}_\sigma$, the analogous operator to P_σ , for the backwards (in time) adjoint state by $(R_\sigma \mathbf{w})_{n,h} = P_h \mathbf{w}(t_{n-1})$ for $1 \leq n \leq N_\tau$. Here P_h denotes the standard L^2 projection on \mathbf{Y}_h already considered in the definition of P_σ . As for the discrete adjoint states, we fix $(R_\sigma \mathbf{w})(t_{n-1}) = (R_\sigma \mathbf{w})_{n,h}$. Analogously to (3.10) and (3.11), we have the estimates for every $\mathbf{w} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$

$$\begin{aligned} & \| \mathbf{w} - R_\sigma \mathbf{w} \|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \\ & \leq C \left\{ \tau \| \mathbf{w}' \|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h^2 \| \mathbf{w} \|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\}, \end{aligned} \quad (3.38)$$

$$\begin{aligned} & \| \mathbf{w} - R_\sigma \mathbf{w} \|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \\ & \leq C \left\{ \frac{\tau}{h} \| \mathbf{w}' \|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \| \mathbf{w} \|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\}. \end{aligned} \quad (3.39)$$

Analogously to the discrete state equation we introduce the notation $\boldsymbol{\eta} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_\sigma = (\boldsymbol{\varphi} - R_\sigma \boldsymbol{\varphi}) + (R_\sigma \boldsymbol{\varphi} - \boldsymbol{\varphi}_\sigma) = \hat{\boldsymbol{\eta}} + \boldsymbol{\eta}_\sigma$. Due to (3.38) it is enough to estimate $\boldsymbol{\eta}_\sigma$ in $L^2(0, T; \mathbf{L}^2(\Omega_h))$. Using [6, Theorem 4.11] and (3.39) we obtain

$$\| \boldsymbol{\eta}_\sigma \|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} + \| \boldsymbol{\eta} \|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \leq Ch. \quad (3.40)$$

We employ a duality argument, similarly to the proof of Theorem 6.

$$\begin{cases} (\mathbf{z}_t, \mathbf{w}) + a(\mathbf{z}, \mathbf{w}) + c(\mathbf{z}, \mathbf{y}_{\mathbf{u}}, \mathbf{w}) + c(\mathbf{y}_{\mathbf{u}}, \mathbf{z}, \mathbf{w}) = (\boldsymbol{\eta}_\sigma, \mathbf{w}), \\ \mathbf{z}_{\mathbf{v}}(0) = 0, \end{cases} \quad (3.41)$$

The discrete version of (3.41) associated to the discrete state $\mathbf{y}_\sigma \approx \mathbf{y}$ is given by

$$\begin{cases} \text{For } n = 1, \dots, N_\tau, \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} \{ a(\mathbf{z}_{n,h}, \mathbf{w}_h) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) \\ + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_h) \} dt = \int_{t_{n-1}}^{t_n} (\boldsymbol{\eta}_{n,h}, \mathbf{w}_h) dt, \\ \mathbf{z}_{0,h} = 0; \end{cases} \quad (3.42)$$

These are the linearized versions of the continuous and discrete Navier-Stokes operators. The following error estimates can be derived easily by using techniques similar to [6, Theorem 4.13].

$$\begin{aligned} & \| \mathbf{z}_\sigma - \mathbf{z} \|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} + \| \mathbf{z}_\sigma - \mathbf{z} \|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \\ & \leq Ch \| \mathbf{z} \|_{\mathbf{H}^{2,1}(\Omega_T)} \leq Ch \| \boldsymbol{\eta}_\sigma \|_{L^2(0, T; \mathbf{L}^2(\Omega_h))}. \end{aligned} \quad (3.43)$$

Subtracting equations (2.9) and (3.35), and taking $\mathbf{w} = \mathbf{w}_h = \mathbf{z}_{n,h}$ we obtain for $\boldsymbol{\eta} = \boldsymbol{\phi} - \boldsymbol{\phi}_\sigma$,

$$\begin{aligned} & (\boldsymbol{\eta}(t_{n-1}) - \boldsymbol{\eta}(t_n), \mathbf{z}_{n,h}) + \int_{t_{n-1}}^{t_n} \{a(\boldsymbol{\eta}, \mathbf{z}_{n,h}) \\ & + c(\mathbf{y}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) + c(\mathbf{z}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t)) \\ & - c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h})\} dt \\ & = \int_{t_{n-1}}^{t_n} (\mathbf{e}(t), \mathbf{z}_{n,h}) dt. \end{aligned}$$

Now, using the decomposition $\boldsymbol{\eta} = \hat{\boldsymbol{\eta}} + \boldsymbol{\eta}_\sigma$ and taking into account $(\hat{\boldsymbol{\eta}}(t_j), \mathbf{w}_h) = 0$ for all $j = 1, \dots, N_\tau$, we infer from the above relation

$$\begin{aligned} & (\boldsymbol{\eta}_{n,h} - \boldsymbol{\eta}_{n+1,h}, \mathbf{z}_{n,h}) + \int_{t_{n-1}}^{t_n} \{a(\boldsymbol{\eta}_{n,h}, \mathbf{z}_{n,h}) \\ & + c(\mathbf{y}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) + c(\mathbf{z}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t)) \\ & - c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h})\} dt \\ & = \int_{t_{n-1}}^{t_n} (\mathbf{e}(t), \mathbf{z}_{n,h}) dt - \int_{t_{n-1}}^{t_n} a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt. \end{aligned} \quad (3.44)$$

Taking $\mathbf{w}_h = \boldsymbol{\eta}_{n,h}$ in (3.42) and subtracting (3.44) to the resulting equation we get,

$$\begin{aligned} & \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}_{n,h}\|_{L^2(\Omega_h)}^2 dt = (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \boldsymbol{\eta}_{n,h}) - (\boldsymbol{\eta}_{n,h} - \boldsymbol{\eta}_{n+1,h}, \mathbf{z}_{n,h}) \\ & + \int_{t_{n-1}}^{t_n} \{c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\eta}_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\eta}_{n,h}) \\ & - c(\mathbf{y}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{z}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t)) \\ & + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\varphi}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h})\} dt \\ & + \int_{t_{n-1}}^{t_n} (\mathbf{e}(t), \mathbf{z}_{n,h}) dt - \int_{t_{n-1}}^{t_n} a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt. \end{aligned}$$

With the identities $\mathbf{e} = \mathbf{y} - \mathbf{y}_\sigma$ and $\boldsymbol{\eta} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_\sigma$, we get

$$\begin{aligned} & \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}_{n,h}\|_{L^2(\Omega_h)}^2 dt = (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \boldsymbol{\eta}_{n,h}) - (\boldsymbol{\eta}_{n,h} - \boldsymbol{\eta}_{n+1,h}, \mathbf{z}_{n,h}) \\ & + \int_{t_{n-1}}^{t_n} \{c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\eta}_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\eta}_{n,h}) \\ & - c(\mathbf{e}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{z}_{n,h}, \mathbf{e}(t), \boldsymbol{\varphi}(t)) \end{aligned}$$

$$\begin{aligned}
& -c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \boldsymbol{\eta}(t)) - c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\eta}(t)) \} dt \\
& + \int_{t_{n-1}}^{t_n} (\mathbf{e}(t), \mathbf{z}_{n,h}) dt - \int_{t_{n-1}}^{t_n} a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt.
\end{aligned}$$

Recalling that $\boldsymbol{\eta}_\sigma - \boldsymbol{\eta} = -\hat{\boldsymbol{\eta}}$, and canceling appropriate terms, the previous identity is reduced to

$$\begin{aligned}
& \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}_{n,h}\|_{L^2(\Omega_h)}^2 dt = (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \boldsymbol{\eta}_{n,h}) - (\boldsymbol{\eta}_{n,h} - \boldsymbol{\eta}_{n+1,h}, \mathbf{z}_{n,h}) \\
& - \int_{t_{n-1}}^{t_n} \{c(\mathbf{e}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) + c(\mathbf{z}_{n,h}, \mathbf{e}(t), \boldsymbol{\varphi}(t)) \\
& + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \hat{\boldsymbol{\eta}}(t)) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \hat{\boldsymbol{\eta}}(t)) \} dt \\
& + \int_{t_{n-1}}^{t_n} (\mathbf{e}(t), \mathbf{z}_{n,h}) dt - \int_{t_{n-1}}^{t_n} a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt.
\end{aligned}$$

Adding the above equalities from $n = N_\tau$ to $n = 1$, and using that $\mathbf{z}_{0,h} = \boldsymbol{\eta}_{N_\tau,h} = 0$, we obtain

$$\begin{aligned}
& \int_0^T \|\boldsymbol{\eta}_{n,h}\|_{L^2(\Omega_h)}^2 dt = - \int_0^T \{c(\mathbf{e}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) + c(\mathbf{z}_{n,h}, \mathbf{e}(t), \boldsymbol{\varphi}(t)) \\
& + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \hat{\boldsymbol{\eta}}(t)) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \hat{\boldsymbol{\eta}}(t)) \} dt \\
& + \int_0^T (\mathbf{e}(t), \mathbf{z}_{n,h}) dt - \int_0^T a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt. \tag{3.45}
\end{aligned}$$

It remains to bound each term on the right hand side. For the first term, note that

$$\begin{aligned}
& \int_0^T c(\mathbf{e}(t), \mathbf{z}_{n,h}, \boldsymbol{\varphi}(t)) dt \\
& \leq C \|\mathbf{e}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \|\mathbf{z}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \|\boldsymbol{\varphi}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \\
& \leq Ch^2 \|\boldsymbol{\eta}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))},
\end{aligned}$$

where we have used the estimate (3.14) and stability estimates of \mathbf{z}_σ and $\boldsymbol{\varphi}$. The second can be handled similarly. For the third one, we proceed as follows:

$$\begin{aligned}
& \int_0^T c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \hat{\boldsymbol{\eta}}(t)) dt = \int_0^T c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h} - \mathbf{z}(t), \hat{\boldsymbol{\eta}}(t)) dt \\
& + \int_0^T c(\mathbf{y}_{n,h}, \mathbf{z}(t), \hat{\boldsymbol{\eta}}(t)) dt \\
& \leq C \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \|\mathbf{z} - \mathbf{z}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \|\hat{\boldsymbol{\eta}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}
\end{aligned}$$

$$\begin{aligned}
& + \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \|\mathbf{z}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \|\hat{\boldsymbol{\eta}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\
& \leq Ch^2 \|\boldsymbol{\eta}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.
\end{aligned}$$

The above inequalities follow from the estimates (3.38), (3.39), (3.43) and stability properties as before. The fourth term of (3.45) is handled similar to the previous estimate, the fifth one is consequence of (3.14) and the stability of \mathbf{z}_σ , while the last one can be treated as follows,

$$\begin{aligned}
\int_0^T a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h}) dt &= \int_0^T a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}_{n,h} - \mathbf{z}(t)) dt + \int_0^T a(\hat{\boldsymbol{\eta}}(t), \mathbf{z}(t)) dt \\
&\leq C(\|\hat{\boldsymbol{\eta}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \|\mathbf{z} - \mathbf{z}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\
&+ \|\hat{\boldsymbol{\eta}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \|\mathbf{z}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}) \leq Ch^2 \|\boldsymbol{\eta}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.
\end{aligned}$$

4 Error estimates of the discrete optimal control problem

As in §3.2, the hypothesis $\tau \leq C_0 h^2$ will be assumed throughout this section. Now, we are ready to prove the main result regarding the error estimates in the $L^2(0, T; \mathbf{L}^2(\Omega))$ norm. First, we formulate the discrete control problem as follows

$$(\mathbf{P}_\sigma) \quad \begin{cases} \min J_\sigma(\mathbf{u}_\sigma) \\ \mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad} \end{cases}$$

$$\begin{aligned}
J_\sigma(\mathbf{u}_\sigma) &= \frac{1}{2} \int_0^T \int_{\Omega_h} |\mathbf{y}_\sigma(\mathbf{u}_\sigma) - \mathbf{y}_d|^2 dx dt \\
&+ \frac{\gamma}{2} \int_{\Omega_h} |\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}|^2 dx + \frac{\lambda}{2} \int_0^T \int_{\Omega_h} |\mathbf{u}_\sigma|^2 dx dt,
\end{aligned}$$

where $\mathbf{y}_{\Omega_h} \in \mathbf{Y}_h$, with $\|\mathbf{y}_\Omega - \mathbf{y}_{\Omega_h}\|_{\mathbf{L}^2(\Omega_h)} \leq Ch$ and $\|\mathbf{y}_{\Omega_h}\|_{\mathbf{H}^1(\Omega_h)} \leq C \forall h > 0$; $\mathbf{y}_\sigma(\mathbf{u}_\sigma)$ is defined in (3.7) with right hand side \mathbf{u}_σ ; and $\mathcal{U}_{\sigma,ad} = \mathcal{U}_{ad} \cap \mathcal{U}_\sigma$. There are three standard ways of choosing \mathcal{U}_σ .

1-Piecewise constant controls:

$$\mathbf{U}_h = \mathbf{U}_{h,0} = \{\mathbf{u}_h \in \mathbf{L}^2(\Omega_h) : \mathbf{u}_h|_K \equiv \mathbf{u}_K \in \mathbb{R}^2 \quad \forall K \in \mathcal{K}_h\}$$

and

$$\mathcal{U}_\sigma = \mathcal{U}_{\sigma,0} = \{\mathbf{u}_\sigma \in L^2(0, T; \mathbf{U}_h) : \mathbf{u}_\sigma|_{(t_{n-1}, t_n]} \in \mathbf{U}_h, \text{ for } 1 \leq n \leq N_\tau\}.$$

2-Piecewise linear controls:

$$\mathbf{U}_h = \mathbf{U}_{h,1} = \{\mathbf{u}_h \in \mathbf{C}(\bar{\Omega}_h) : \mathbf{u}_h|_K \in \mathcal{P}_1(K)^2 \quad \forall K \in \mathcal{K}_h\}$$

and

$$\mathcal{U}_\sigma = \mathcal{U}_{\sigma,1} = \{\mathbf{u}_\sigma \in L^2(0, T; \mathbf{U}_h) : \mathbf{u}_\sigma|_{(t_{n-1}, t_n)} \in \mathbf{U}_h \text{ for } 1 \leq n \leq N_\tau\}.$$

3-Variational discretization: $\mathbf{U}_h = \mathbf{U}_{h,2} = \mathbf{L}^2(\Omega_h)$ and $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,2} = L^2(0, T; \mathbf{L}^2(\Omega_h))$.

It is easy to check that for any of the three precedent choices, the discrete problem (P_σ) has at least one solution. The next theorem states that the family of problems (P_σ) realizes a good approximation of problem (P). We refer the reader to [6, Theorems 4.13 and 4.15] for a detailed proof.

Theorem 8 *For every $\sigma = (\tau, h)$ let $\bar{\mathbf{u}}_\sigma$ be a global solution of problem (P_σ) , then the sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ is bounded in $L^2(0, T; \mathbf{L}^2(\Omega))$ and there exist subsequences, denoted in the same way, converging to a point $\bar{\mathbf{u}}$ weakly in $L^2(0, T; \mathbf{L}^2(\Omega))$. Any of these limit points is a solution of problem (P). Moreover, we have*

$$\lim_{\sigma \rightarrow 0} \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} = 0 \quad \text{and} \quad \lim_{\sigma \rightarrow 0} J_\sigma(\bar{\mathbf{u}}_\sigma) = J(\bar{\mathbf{u}}). \quad (4.1)$$

Conversely, let $\bar{\mathbf{u}}$ be a strict local minimum of (P), then there exists a sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ of local minima of problems (P_σ) such that (4.1) holds.

The key ingredient of the proof is the use of the projection of any solution $\tilde{\mathbf{u}}$ of problem (P) on $\mathcal{U}_{\sigma,ad}$. Then, if \mathbf{u}_σ is the $L^2(0, T; \mathbf{L}^2(\Omega_h))$ projection of $\tilde{\mathbf{u}}$ on $\mathcal{U}_{\sigma,ad}$, then $\|\tilde{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0$ when $\sigma \rightarrow 0$. In the case $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,0}$, then \mathbf{u}_σ is given by

$$\mathbf{u}_\sigma = \sum_{n=1}^{N_\tau} \sum_{K \in \mathcal{K}_h} \mathbf{u}_{n,K} \chi_n \chi_K, \quad \text{with} \quad \mathbf{u}_{n,K} = \frac{1}{|K| \tau_n} \int_{t_{n-1}}^{t_n} \int_K \tilde{\mathbf{u}}(t, x) \, dx dt,$$

where χ_n, χ_K denote the characteristic functions on $(t_{n-1}, t_n]$ and on the element T respectively. Then, \mathbf{u}_σ is the $L^2(0, T; \mathbf{L}^2(\Omega_h))$ projection of $\tilde{\mathbf{u}}$ on \mathcal{U}_σ .

If $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,2}$, then $\mathbf{u}_\sigma = \tilde{\mathbf{u}}$. Though we do not have an explicit expression for the projection in the case $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,1}$, it is not difficult to prove that such a projection satisfies $\|\tilde{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0$. Indeed, it is enough to use that

$$\|\tilde{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \leq \|\tilde{\mathbf{u}} - \mathbf{I}_\sigma \tilde{\mathbf{u}}\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0,$$

where

$$\mathbf{I}_\sigma \tilde{\mathbf{u}} = \sum_{n=1}^{N_\tau} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \mathbf{I}_h \tilde{\mathbf{u}}(t) \, dt \, \chi_n,$$

$\mathbf{I}_h : C(\bar{\Omega})^2 \rightarrow \mathbf{U}_{h,1}$ denoting the standard linear interpolation operator. Recall that Theorem (3) states that any local solution $\tilde{\mathbf{u}}$ of (P) belongs to $L^2(0, T; \mathbf{W}^{1,p}(\Omega)) \subset L^2(0, T; C(\bar{\Omega})^2)$ for $p > 2$.

Theorem 8 can be also deduced by using the abstract result [9, Theorem 2.11]. Indeed, with the notation of [9], we can take $\mathcal{A} = U_2 = U_\infty = L^2(0, T; \mathbf{L}^2(\Omega))$ and the assumptions (A1) and (A3)-(A6) of [9] can be easily proved by using the previous analysis carried out in the sections §2 and §3.

Hereafter, $\tilde{\mathbf{u}}$ will denote a local solution of (P) and for every σ , $\tilde{\mathbf{u}}_\sigma$ denotes a local solution of (P_σ) such that $\|\tilde{\mathbf{u}} - \tilde{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0$; see Theorem 8. We also denote by $\tilde{\mathbf{y}}$ and $\tilde{\varphi}$ the state and adjoint associated to $\tilde{\mathbf{u}}$, and $\tilde{\mathbf{y}}_\sigma$ and $\tilde{\varphi}_\sigma$ will denote the discrete state and adjoint state corresponding to $\tilde{\mathbf{u}}_\sigma$. The goal is to estimate the rate of the convergence $(\tilde{\mathbf{u}}_\sigma, \tilde{\mathbf{y}}_\sigma, \tilde{\varphi}_\sigma) \rightarrow (\tilde{\mathbf{u}}, \tilde{\mathbf{y}}, \tilde{\varphi})$. As in [6, Section 4] and [9], all the elements $\mathbf{u}_\sigma \in \mathcal{U}_\sigma$, for $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,0}$ and $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,1}$, are extended to $(0, T) \times \Omega$ by setting $\mathbf{u}_\sigma(t, x) = \tilde{\mathbf{u}}(t, x)$ for $(t, x) \in (0, T) \times (\Omega \setminus \Omega_h)$.

Let us write $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2)$. Associated to the components \tilde{u}_j , $j = 1, 2$, for every $t \in (0, T)$, and motivated by (2.19)-(2.20), we split the elements $I_n \times K$, with $I_n = (t_{n-1}, t_n]$ and $K \in \mathcal{K}_h$, as follows: $\mathcal{T}_\sigma = \mathcal{T}_{\sigma,1}^j \cup \mathcal{T}_{\sigma,2}^j \cup \mathcal{T}_{\sigma,3}^j$, $j = 1, 2$, where

$$\mathcal{T}_\sigma = \{I_n \times K : 1 \leq n \leq N_\tau \text{ and } K \in \mathcal{K}_h\},$$

$$\mathcal{T}_{\sigma,1}^j = \{I_n \times K \in \mathcal{T}_\sigma : \tilde{\varphi}_j(t, x) + \lambda \tilde{u}_j(t, x) \neq 0 \forall (t, x) \in I_n \times K\},$$

$$\mathcal{T}_{\sigma,2}^j = \{I_n \times K \in \mathcal{T}_\sigma : \tilde{\varphi}_j(t, x) + \lambda \tilde{u}_j(t, x) = 0 \forall (t, x) \in I_n \times K\},$$

$$\mathcal{T}_{\sigma,3}^j = \mathcal{T}_\sigma \setminus (\mathcal{T}_{\sigma,1}^j \cup \mathcal{T}_{\sigma,2}^j).$$

Finally, let us denote

$$E_\sigma = \|\tilde{\mathbf{u}} - \tilde{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} + \|\tilde{\mathbf{y}} - \tilde{\mathbf{y}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} + \|\tilde{\varphi} - \tilde{\varphi}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))}, \quad (4.2)$$

$$\mathcal{E}_\sigma = \|\tilde{\mathbf{y}} - \tilde{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} + \|\tilde{\mathbf{y}} - \tilde{\mathbf{y}}_\sigma\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} + \|\tilde{\varphi} - \tilde{\varphi}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} + \|\tilde{\varphi} - \tilde{\varphi}_\sigma\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))}. \quad (4.3)$$

Then we have the following error estimates.

Theorem 9 *Suppose that (2.22) holds. Moreover, if $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,1}$ we also assume that $\mathbf{y}_d \in \mathbf{L}^p(\Omega_T)$, $\mathbf{y}_\Omega \in \mathbf{Y} \cap \mathbf{W}^{1,p}(\Omega)$, with $3 < p < +\infty$, Γ is of class C^3 , and for some constant $M > 0$*

$$\sum_{j=1}^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} |K| \tau_n \leq Mh. \quad (4.4)$$

Then, then we have the following estimates

$$E_\sigma \leq \begin{cases} Ch & \text{if } \mathcal{U}_\sigma = \mathcal{U}_{\sigma,0}, \\ C_p h^{\frac{3}{2} - \frac{2}{p}} & \text{if } \mathcal{U}_\sigma = \mathcal{U}_{\sigma,1}, \\ Ch^2 & \text{if } \mathcal{U}_\sigma = \mathcal{U}_{\sigma,2}, \end{cases} \quad (4.5)$$

$$\mathcal{E}_\sigma \leq Ch \text{ in all cases.} \quad (4.6)$$

Proof First we observe that it is enough to prove the estimates for $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}$, the others are consequence of Corollary 1 and Theorem 7. To prove this estimate, we will use [9, Theorem 2.14]. It is immediate to verify that the assumptions of [9, Theorem 2.14] hold. From (2.7), (3.36) and (3.37) we deduce

$$\begin{aligned} & |[J'_\sigma(\mathbf{u}) - J'(\mathbf{u})](\mathbf{u}_\sigma - \bar{\mathbf{u}})| \\ & \leq Ch^2 \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \quad \forall (\mathbf{u}, \mathbf{u}_\sigma) \in \mathcal{U}_{ad} \times \mathcal{U}_{\sigma,ad}. \end{aligned}$$

This means that $\varepsilon_h = Ch^2$ in assumption (A7) of [9]. Hence, the inequality (2.19) of [9, Theorem 2.14] can be written

$$\begin{aligned} & \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\ & \leq C[h^4 + \|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}})]^{1/2} \quad (4.7) \end{aligned}$$

for every $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad}$ and for all $|\sigma| < |\sigma_0|$. Now, we have to choose \mathbf{u}_σ conveniently to obtain (4.5). To this end, we will distinguish the three different choices for \mathcal{U}_σ .

Case 1: $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,0}$. Here we take \mathbf{u}_σ as the \mathbf{L}^2 -projection of $\bar{\mathbf{u}}$ on \mathcal{U}_σ , which is given by

$$\mathbf{u}_\sigma = \sum_{n=1}^{N_\tau} \sum_{K \in \mathcal{K}_h} \mathbf{u}_{n,K} \chi_n \chi_K, \quad \text{with } \mathbf{u}_{n,K} = \frac{1}{|K| \tau_n} \int_{t_{n-1}}^{t_n} \int_K \bar{\mathbf{u}}(t, x) \, dx dt.$$

Then, we have that $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad}$ and

$$\|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) \leq Ch^2;$$

see [6, Lemma 4.17 and (4.45)] for the proof. It is enough to insert this inequality in (4.7) to obtain the desired estimate.

Case 2: $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,1}$. Now we take

$$\mathbf{u}_\sigma = \mathbf{I}_\sigma \bar{\mathbf{u}} = \sum_{n=1}^{N_\tau} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \mathbf{I}_h \bar{\mathbf{u}}(t) \, dt \chi_n,$$

$\mathbf{I}_h : C(\bar{\Omega})^2 \rightarrow \mathbf{U}_{h,1}$ denoting the standard linear interpolation operator. Again we have that $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad}$, and from Lemma 3 below, we deduce

$$\|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) \leq Ch^{3-4/p}.$$

Once again this inequality and (4.7) proves the estimate.

Case 3: $\mathcal{U}_\sigma = \mathcal{U}_{\sigma,2}$. It is enough to take $\mathbf{u}_\sigma = \bar{\mathbf{u}}$ in (4.7) to deduce the estimate of order h^2 .

Lemma 3 *Let us assume that $\mathbf{y}_d \in \mathbf{L}^p(\Omega_T)$, $\mathbf{y}_\Omega \in \mathbf{Y} \cap \mathbf{W}^{1,p}(\Omega)$, with $3 < p < +\infty$, Γ is of class C^3 , and (4.4) holds. Let $\mathbf{u}_\sigma = \mathbf{I}_\sigma \bar{\mathbf{u}}$, then we have the estimates*

$$\|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq Ch^{\frac{3}{2}-\frac{1}{p}}, \quad (4.8)$$

$$J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) \leq Ch^{3-\frac{4}{p}}. \quad (4.9)$$

Proof The proof is divided into two parts.

1 - *Proof of the estimate (4.8).*

$$\begin{aligned} & \|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\ &= \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \left\| \bar{\mathbf{u}}(t) - \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \mathbf{I}_h \bar{\mathbf{u}}(s) ds \right\|_{\mathbf{L}^2(\Omega_h)}^2 dt \right\}^{1/2} \\ &\leq \left\{ \sum_{n=1}^{N_\tau} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \int_{t_{n-1}}^{t_n} \|\bar{\mathbf{u}}(t) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds dt \right\}^{1/2} \\ &\leq \left\{ \sum_{n=1}^{N_\tau} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \int_{t_{n-1}}^{t_n} \|\bar{\mathbf{u}}(t) - \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds dt \right\}^{1/2} \\ &\quad + \left\{ \sum_{n=1}^{N_\tau} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \int_{t_{n-1}}^{t_n} \|\bar{\mathbf{u}}(s) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds dt \right\}^{1/2} \\ &\leq \tau \left\{ \int_0^T \|\bar{\mathbf{u}}'(\theta)\|_{\mathbf{L}^2(\Omega_h)}^2 d\theta \right\}^{1/2} + \left\{ \int_0^T \|\bar{\mathbf{u}}(s) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds \right\}^{1/2} \\ &\leq Ch^2 \|\bar{\mathbf{u}}'\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \left\{ \int_0^T \|\bar{\mathbf{u}}(s) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds \right\}^{1/2}. \end{aligned}$$

Let us estimate the last term. First we write

$$\begin{aligned} & \left\{ \int_0^T \|\bar{\mathbf{u}}(s) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds \right\}^{1/2} \\ & \leq \sum_{j=1}^2 \left\{ \sum_{n=1}^{N_\tau} \sum_{K \in \mathcal{T}_h} \int_{t_{n-1}}^{t_n} \|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^2(K)}^2 dt \right\}^{1/2}. \end{aligned}$$

Now, from (2.12) we deduce that $\bar{u}_j \equiv \alpha_j$ or $\bar{u}_j \equiv \beta_j$ in $I_n \times K$ for every $I_n \times K \in \mathcal{T}_{\sigma,1}^j$, therefore $\|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^2(K)}^2 = 0$ for every $t \in I_n$.

To deal with the other two components of \mathcal{T}_σ we use the regularity assumptions. From Solonnikov's Theorem [39], see also [19], we deduce that $\bar{\varphi} \in L^p(0, T; \mathbf{W}^{2,p}(\Omega)) \subset L^p(0, T; \mathbf{C}^1(\bar{\Omega}))$ and $\bar{\varphi}' \in \mathbf{L}^p(\Omega_T)$, hence (2.12) implies that $\bar{\mathbf{u}} \in \mathbf{W}^{1,p}(\Omega_T)$.

Let us consider the case $I_n \times K \in \mathcal{T}_{\sigma,2}^j$, then $\bar{\varphi}_j(t, x) + \lambda \bar{u}_j(t, x) = 0$ in $I_n \times K$. Therefore, we have that $\bar{u}_j \in L^p(0, T; \mathbf{W}^{2,p}(K))$, hence

$$\begin{aligned} & \sum_{I_n \times K \in \mathcal{T}_{\sigma,2}^j} \int_{t_{n-1}}^{t_n} \|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^2(K)}^2 dt \\ & \leq \sum_{I_n \times K \in \mathcal{T}_{\sigma,2}^j} \tau_n^{\frac{p-2}{p}} |K|^{\frac{p-2}{p}} \left(\int_{t_{n-1}}^{t_n} \|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^p(K)}^p dt \right)^{\frac{2}{p}} \\ & \leq Ch^4 \sum_{I_n \times K \in \mathcal{T}_{\sigma,2}^j} \tau_n^{\frac{p-2}{p}} |K|^{\frac{p-2}{p}} \left(\int_{t_{n-1}}^{t_n} \|\bar{u}_j\|_{\mathbf{W}^{2,p}(K)}^p dt \right)^{\frac{2}{p}} \\ & = \frac{C}{\lambda^2} h^4 \sum_{I_n \times K \in \mathcal{T}_{\sigma,2}^j} \tau_n^{\frac{p-2}{p}} |K|^{\frac{p-2}{p}} \left(\int_{t_{n-1}}^{t_n} \|\bar{\varphi}_j\|_{\mathbf{W}^{2,p}(K)}^p dt \right)^{\frac{2}{p}} \\ & \leq Ch^4 \left(\sum_{I_n \times K \in \mathcal{T}_{\sigma,2}^j} |K| \tau_n \right)^{\frac{p-2}{p}} \|\bar{\varphi}_j\|_{L^p(0,T; \mathbf{W}^{2,p}(\Omega))}^2 \\ & \leq C(|\Omega|T)^{\frac{p-2}{p}} h^4 \|\bar{\varphi}_j\|_{L^p(0,T; \mathbf{W}^{2,p}(\Omega))}^2. \end{aligned}$$

Finally, we consider the case where $I_n \times K \in \mathcal{T}_{\sigma,3}^j$. Now, we have

$$\sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^2(K)}^2 dt$$

$$\begin{aligned}
&\leq \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} (|K|\tau_n)^{\frac{p-2}{p}} \left(\int_{t_{n-1}}^{t_n} \|\bar{u}_j(t) - I_h \bar{u}_j(t)\|_{\mathbf{L}^p(K)}^p dt \right)^{\frac{2}{p}} \\
&\leq Ch^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} (|K|\tau_n)^{\frac{p-2}{p}} \left(\int_{t_{n-1}}^{t_n} \|\bar{u}_j(t)\|_{\mathbf{W}^{1,p}(K)}^p dt \right)^{\frac{2}{p}} \\
&\leq Ch^2 \left(\sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} |K|\tau_n \right)^{\frac{p-2}{p}} \|\bar{u}_j\|_{L^p(0,T;\mathbf{W}^{1,p}(\Omega))}^2 \\
&\leq Ch^{3-\frac{2}{p}} \|\bar{u}_j\|_{L^p(0,T;\mathbf{W}^{1,p}(\Omega))}^2,
\end{aligned}$$

where we have used (4.4).

The last arguments lead to

$$\begin{aligned}
&\left\{ \int_0^T \|\bar{\mathbf{u}}(s) - \mathbf{I}_h \bar{\mathbf{u}}(s)\|_{\mathbf{L}^2(\Omega_h)}^2 ds \right\}^{1/2} \\
&\leq Ch^{\frac{3}{2}-\frac{1}{p}} \left(\|\bar{\varphi}\|_{L^p(0,T;\mathbf{W}^{2,p}(\Omega))} + \|\bar{\mathbf{u}}\|_{L^p(0,T;\mathbf{W}^{1,p}(\Omega))} \right),
\end{aligned}$$

which concludes the proof of (4.8).

2 - *Proof of the estimate (4.9).* Let us set $\bar{\mathbf{d}} = \bar{\varphi} + \lambda \bar{\mathbf{u}}$ and $\bar{\mathbf{d}} = (\bar{d}_1, \bar{d}_2)$. Recalling that, as every element of \mathcal{U}_σ , \mathbf{u}_σ is extended to $\Omega \setminus \Omega_h$ by $\mathbf{u}_\sigma(t, x) = \bar{\mathbf{u}}(t, x)$ for $x \in \Omega \setminus \Omega_h$, we have

$$J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) = \sum_{j=1}^2 \sum_{n=1}^{N_\tau} \sum_{K \in \mathcal{T}_h} \int_{t_{n-1}}^{t_n} \int_K \bar{d}_j(u_j - \bar{u}_j) dx dt.$$

This expression can be simplified as follows. As we already mentioned in the proof of the previous lemma, if $I_n \times K \in \mathcal{T}_{\sigma,1}^j$, then $u_j(t, x) = \bar{u}_j(t, x)$ for all $(t, x) \in I_n \times K$, both of them being either α_j or β_j at the same time. Therefore, the integral in $I_n \times K$ is zero. If $I_n \times K \in \mathcal{T}_{\sigma,2}^j$, then $\bar{d}_j(t, x) = 0$ for every $(t, x) \in I_n \times K$. Hence, the integral in $I_n \times K$ is again zero. Finally, if $I_n \times K \in \mathcal{T}_{\sigma,3}^j$, then $\bar{d}_j \neq 0$ in $I_n \times K$, but there exists at least one point $(s_n^j, \xi_K^j) \in I_n \times K$ such that $\bar{d}_j(s_n^j, \xi_K^j) = 0$. Hence, we have

$$\begin{aligned}
&J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) = \\
&\sum_{j=1}^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \int_K (\bar{d}_j(t, x) - \bar{d}_j(s_n^j, \xi_K^j))(u_j(t, x) - \bar{u}_j(t, x)) dx dt
\end{aligned}$$

$$\begin{aligned}
&\leq \left\{ \sum_{j=1}^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \int_K (\bar{d}_j(t, x) - \bar{d}_j(s_n^j, \xi_K^j))^2 dx dt \right\}^{\frac{1}{2}} \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{\mathbf{L}^2(\Omega_T)} \\
&\leq Ch^{\frac{3}{2} - \frac{1}{p}} \left\{ \sum_{j=1}^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \int_K (\bar{d}_j(t, x) - \bar{d}_j(s_n^j, \xi_K^j))^2 dx dt \right\}^{\frac{1}{2}}, \\
\end{aligned} \tag{4.10}$$

the last inequality being an immediate consequence of (4.8). Now, we use that $\bar{\mathbf{d}} \in L^p(0, T; \mathbf{W}^{2,p}(\Omega))$ and, due to $p > 3$, $\bar{\mathbf{d}} \in \mathbf{W}^{1,p}(\Omega_T) \subset \mathbf{C}^{0,\eta}(\bar{\Omega}_T)$ for $\eta = 1 - \frac{3}{p}$; see, for instance, Nečas [34]. Using this properties, we get

$$\begin{aligned}
&\left\{ \sum_{j=1}^2 \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \int_K (\bar{d}_j(t, x) - \bar{d}_j(s_n^j, \xi_K^j))^2 dx dt \right\}^{\frac{1}{2}} \\
&\leq Ch^{1 - \frac{3}{p}} \left\{ \sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} \int_{t_{n-1}}^{t_n} \int_K \|\bar{\mathbf{d}}\|_{\mathbf{W}^{1,p}(\Omega_T)}^2 dx dt \right\}^{\frac{1}{2}} \\
&\leq Ch^{1 - \frac{3}{p}} \left(\sum_{I_n \times K \in \mathcal{T}_{\sigma,3}^j} |K| \tau_n \right)^{\frac{1}{2}} \|\bar{\mathbf{d}}\|_{\mathbf{W}^{1,p}(\Omega_T)} \leq Ch^{\frac{3}{2} - \frac{3}{p}} \|\bar{\mathbf{d}}\|_{\mathbf{W}^{1,p}(\Omega_T)}.
\end{aligned}$$

This estimate along with (4.10) lead to (4.9).

References

1. F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynamics*, 1:303–325, 1990.
2. T. Apel and T. Flaig, Crank-Nicolson schemes for optimal control problems with evolution equations. *SIAM J. Numer. Anal.*, 50(3): 1484–1512, 2012.
3. J.F. Bonnans and H. Zidani. Optimal control problems with partially polyhedral constraints. *SIAM J. Control Optim.*, 37(6):1726–1741, 1999.
4. S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, Berlin, Heidelberg, 1994.
5. E. Casas. An optimal control problem governed by the evolution Navier-Stokes equations. In S.S. Sritharan, editor, *Optimal Control of Viscous Flows*, Philadelphia, 1998. Frontiers in Applied Mathematics, SIAM.
6. E. Casas and K. Chrysafinos. A discontinuous Galerkin time-stepping scheme for the velocity tracking problem. *SIAM J. Numer. Anal.*, 50(5):2281–2306, 2012.

7. E. Casas and M. Mateos. Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints. *SIAM J. on Control & Optim.*, 40(5):1431–1454, 2002.
8. E. Casas, M. Mateos, and J.-P. Raymond. Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier-Stokes equations. *SIAM J. on Control & Optim.*, 46(3):952–982, 2007.
9. E. Casas, and F. Tröltzsch. A general theorem on error estimates with application to elliptic optimal control problems. *Comput. Optim. Appl.*, DOI 10.1007/s10589-011-9453-8, 2012.
10. E. Casas and J.-P. Raymond. Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. *SIAM J. on Control & Optim.*, 45(5):1586–1611, 2006.
11. K. Chrysafinos. Analysis and finite element approximations for distributed optimal control problems for implicit parabolic PDE's. *J. Comput. Appl. Math.*, 231:327-348, 2009.
12. K. Chrysafinos. Convergence of discontinuous Galerkin approximations of an optimal control problem associated to semilinear parabolic pde's. *ESAIM M²AN*, 44(1):189–206, 2010.
13. K. Chrysafinos and N.J. Walkington. Discontinuous Galerkin approximations of the Stokes and Navier-Stokes equations. *Math. Comput.*, 79(272):2135–2167, 2010.
14. K. Deckelnick and M. Hinze. Semidiscretization and error estimates for distributed control of the instationary Navier-Stokes equations. *Numer. Math.*, 97:297–320, 2000.
15. K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.
16. K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $l_\infty(l^2)$ and $l_\infty(l_\infty)$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.
17. K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems IV: Nonlinear problems. *SIAM J. Numer. Anal.*, 32(6):1729–1749, 1995.
18. D. Estep and S. Larsson. The discontinuous Galerkin method for semilinear parabolic equations. *RAIRO Modél. Math. Anal. Numér.*, 27:35–54, 1993.
19. M. Geissert, M. Hess, M. Hieber, C. Schwarz and K. Stavrakidis. Maximal L^p - L^q -estimates for the Stokes equation: a short proof of Solonnikov's theorem. *J. Math. Fluid Mech.*, 12:47–60, 2010.
20. P. Girault and P.A. Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*. Springer-Verlag, Berlin, Heidelberg, New York, Tokio, 1986.
21. M.D. Gunzburger. *Perspectives in flow control and optimization*. SIAM, Advances in Design and Control, Philadelphia, 2003.
22. M.D. Gunzburger. *Flow Control*. Springer-Verlag, New York, 1995.
23. M.D. Gunzburger and S. Manservigi. The velocity tracking problem for Navier-Stokes flows with bounded distributed control. *SIAM J. Control Optim.*, 37(6):1913–1945, 1999.
24. M.D. Gunzburger and S. Manservigi. Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control. *SIAM J. Numer. Anal.*, 37:1481–1512, 2000.

25. M. Hinze. A Variational discretization concept in control constrained optimization: The linear quadratic case. *Comput. Optim. and Appl.*, 30:45–61, 2005.
26. M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM J. Control Optim.*, 40(3):925–946, 2001.
27. O.A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow*. Gordon and Breach, New York, second edition edition, 1969. English translation.
28. J.L. Lions. *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*. Dunod, Paris, 1969.
29. J.L. Lions and E. Magenes. *Problèmes aux Limites non Homogènes*. Dunod, Paris, 1968.
30. C. Meyer and A. Rösch. Superconvergence properties of optimal control problems. *SIAM J. Control and Optim.*, 43: 353–376, 2004.
31. D. Meidner and B. Vexler. A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part I: Problems without control constraints. *SIAM J. Control and Optim.*, 47(3):1150–1177, 2008.
32. D. Meidner and B. Vexler. A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part II: Problems with control constraints. *SIAM J. Control and Optim.*, 47(3):1301–1329, 2008.
33. D. Meidner and B. Vexler. A-priori error analysis of the Petrov-Galerkin Crank-Nicolson scheme for parabolic optimal control problems. *SIAM J. Control and Optim.*, 49(5), 2183–2211, 2011.
34. J. Nečas. *Les Méthodes Directes en Théorie des Equations Elliptiques*. Editeurs Academia, Prague, 1967.
35. I. Neitzel and B. Vexler. A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems. *Numer. Math.*, 120:345–386, 2012.
36. P.A. Raviart and J.M. Thomas. *Introduction à L'analyse Numérique des Equations aux Dérivées Partielles*. Masson, Paris, 1983.
37. A. Rösch. Error estimates for parabolic optimal control problems with control constraints. *Z. Anal. Anwendungen*, 23:353–376, 2004.
38. A. Rösch and B. Vexler. Optimal control of the Stokes equations: A priori error analysis for finite element discretization with postprocessing. *SIAM J. Numer. Anal.*, 44:1903–1920, 2006.
39. V.A. Solonnikov. Estimates for solutions of nonstationary Navier-Stokes equations. *J. Soviet. Math.*, 8:213–317, 1977.
40. S.S. Sritharan. *Optimal control of viscous flow*. SIAM, Philadelphia, 1998.
41. R. Temam. *Navier-Stokes Equations*. North-Holland, Amsterdam, 1979.
42. V. Thomée. *Galerkin finite element methods for parabolic problems*. Springer-Verlag, Berlin, 1997.
43. F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Philadelphia, 2010.
44. F. Tröltzsch and D. Wachsmuth. Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations. *ESAIM: COCV*, 12:93–119, 2006.